



Effects of Task Conditions on Spoken Performance in Retelling

著者	Yokoi Yuichiro
内容記述	この博士論文は内容の要約のみの公開（または一部非公開）になっています
year	2018
その他のタイトル	再話におけるタスク条件が発話にもたらす影響
学位授与大学	筑波大学 (University of Tsukuba)
学位授与年度	2018
報告番号	12102甲第8786号
URL	http://hdl.handle.net/2241/00153779

Effects of Task Conditions on Spoken Performance in Retelling

(再話におけるタスク条件が発話にもたらす影響)

筑波大学大学院 人文社会科学研究所

現代語・現代文化専攻 言語情報分野

201230048 横内 裕一郎

博士論文要約

英語教育において、大学入試改革の計画や次期学習指導要領が公表されたことにより、スピーキングの指導と4技能を統合した指導や評価への関心とニーズがますます高まっている。スピーキングの評価は4技能の中でも最も難しく (Fulcher, 2003)、試験の実施や採点に時間がかかることや入試等でスピーキングテストが課されてこなかったこと (根岸, 2013) などが、これまでスピーキング評価の実施が教育現場で避けられてきた原因である。

スピーキングテストで使用されることが想定されるタスクの中でも、再話課題は問題作成や問題内容の調整が容易で、刺激文に特定の表現や文法を入れ込むことが容易であるため、スピーキング能力を評価するだけでなく、既習事項の表現を使えるかどうかを確認することにも使用することができる (平井, 2015)。再話は情報が与えられた上で内容にあわせて発話を行う活動であることから、典型的な技能統合型のタスクであると言える。技能統合型のタスクの利点は、文章を読んだり聞いたりしてその内容を元に自身の発表につなげることから、English for academic purposeの観点で真正性があることと、受験者に共通の情報が与えられるため、背景知識が発話に与える影響が制限されることから公平性が高いことが挙げられる。さらに、受験者が与えられた情報から内容と表現を学ぶことができるため、学習への高い波及効果が期待される(e.g., Barkaoui et al., 2013; Cumming, 2013; Huang, Hung, & Hong, 2016; Plakans, 2007)。投野他 (2013) によるCEFR-Jのspoken production B1.2には「短い読み物か短い

新聞記事であれば、ある程度の流暢さをもって、自分の感想や考えを加えながら、あらすじや要点を順序だてて伝えることができる」と記述がある。CEFRのB1レベルは英検2級程度とみなされることから、再話課題において十分に能力が発揮できれば、高校卒業者に求められるスピーキング能力を満たしていると考えられるだろう。以上のように、スピーキング能力を評価するためのタスクとして再話は有効な活動であるものの、再話課題における刺激文の提示方法や長さ、難易度をはじめとするタスクの特徴が発話に与える影響はまだ精査がなされていない。

本博士論文は4つの研究で構成されており、それぞれの研究にパイロットスタディを含め、7つの実験・分析が実施された。研究1では再話と要約課題におけるスピーキングパフォーマンスの違い (pilot study) と指示の与え方による発話への影響を調査し、研究2では刺激文の長さ、文章の難易度、提示方法 (リーディング・リスニング) の差が発話に与える影響を調査した。続いて、再話課題の前に準備時間を与えた場合と音読課題を与えた場合、つまり、本タスクの前にプレタスクを置いた場合にどのようにパフォーマンスが変化するかを調査した (研究3)。研究1から3では、発話の複雑さ・正確さ・流暢さ (CAF) の観点、及び、Hirai and Koizumi (2008, 2013)による empirically derived, binary-choice, boundary-definition (EBB) scale による評価結果を観察し、異なる条件下で発話にどのような影響が現れるかを検証した。その上で、これらの結果を統合するため、研究4において多相ラッシュモデルによる分析を実施し、研究1〜3で用いられたタスク及び刺激文の難易度を明らかにした。

研究1では、パイロットスタディにおいて日本人大学生29名を対象に「再話」と「要約」の両課題における発話の質に違いが生じるかどうかを検証した。また、受験者と本実験に参加した評価者が、それぞれ再話と要約をどのような活動と捉えているかをインタビューを通じて調査した。

RQ1A-1: Do spoken performances differ between retelling tasks and summarizing tasks?

RQ1A-2: How do the test takers and raters of this experiment interpret retelling and summarizing tasks?

協力者は日本人大学生 29 名で、そのうち 21 名のデータを分析対象とした。2 種類の刺激文を用い、再話と要約活動を実験協力者に課し、その発話を比較した。その結果、本実験で使用したすべての指標 (number of token words, number of word types, number of syllables, number of 1 second or longer silent pauses, number of 1 second or longer silent pauses, total length of silent pauses, words per minute, syllables per second, number of disfluency markers, errors per AS-unit, Guiraud index, type token ratio, words per AS-unit) で差は生じなかった。RQ1A-2 について、実験協力者と評価者に再話と要約のそれぞれのタスクをどのような課題と捉えていたかをインタビューした結果、実験協力者からは 11 名の回答が得られ、そのうち半数以上 6 名が再話と要約が同質なものであるという回答があった。一方、要約のほうが簡単だと感じた回答が 3 件あり、回答の内訳は「要約のほうが自分で文章を構築できるため」という回答が得られた。また、再話の方が簡単であると判断した回答が 2 件寄せられ、こちらでは「記憶に残った文章をそのまま使えるため簡単だと感じた」という意見があった。評価者からの意見としては、話の時系列に沿って発話が展開していれば再話、特に、刺激文と同じ表現が繰り返されている場合にそうだと判定し、同じ表現が多用されている場合には再話、言い換えが多い場合には要約であるという意見が寄せられた。この結果から、再話と要約のタスクは少なくとも発話の CAF の観点で発話パフォーマンスが変わらないことと、実験協力者と評価者へのインタビューの結果から、基本的に再話と要約は似た課題であると結論づけた。

続いて、パイロットスタディの結果を受け、タスク前の指示の違いがパフォーマンスに影響を及ぼすかを調査することを目的に、それぞれのタスクに明確な指示を与えたうえでデータ収集を行った。それに加え、本研究では、刺激文の長さで再話・要約課題の関係を調査し、さらにどちらの課題で刺激文と同じ表現が多用されるかについても調査を行った。なお、本研究における「同じ表現」は本文と配列が同じ 3 語で構成される表現を指す。以下が本研究のリサーチエスチョン (RQs) である。

RQ1B-1: Do the performances in summarizing tasks differ from the performances in retelling tasks in terms of CAF and raters' evaluations?

RQ1B-2: Does text length affect oral performance in retelling and summarizing tasks?

RQ1B-3: Do participants use as the original expressions in retelling tasks more frequently than in summarizing tasks?

日本人大学生 46 名を対象にデータ収集を行い、そのうち 37 名分のデータが分析に使用された。刺激文の表現を積極的に使用するよう指示を与えた場合 (A) と、刺激文の表現を出来る限り言い換えるよう指示を与えた場合 (B) において CAF の観点から発話の質を観察した。実験の結果、再話と要約の両課題は、CAF の観点において発話の質に差がなく、異なる指示が与えられた場合でも同様の結果がもたらされた。一方、刺激文で使用された表現と同様の表現がどの程度使用されたかを検証した結果、有意差は見られなかったものの、(A) の群で刺激文中の表現と同じ表現を多く使用している可能性があることが示唆された。

研究 2 は実験 2A と 2B で構成されており、刺激文の長さ・難易度・提示方法が発話に与える影響を明らかにすることを目的に実験を実施した。実験 2A では、刺激文の長さ・難易度・提示方法の 3 要素が発話に影響を与えるかを調査した。

RQ2A-1: How does the length of the input material affect the utterances in retelling tasks?

RQ2A-2: How does the difficulty of the input material affect the utterance in retelling tasks?

RQ2A-3: Does the input mode affect the utterance in retelling tasks?

実験 2A には、日本人大学生 59 名が参加し、そのうち 56 名が分析対象となった。4 種類の文章を用いて再話課題を行った結果、与えられる文章が長い場合、発話量が多い上に語彙的多様性や流暢さの指標で短い文章が刺激として与えられた場合よりも高い値を示した。一方、文章の難易度が発話に与える影響を調査した結果、Flesch-Kincaid Grade Level が低い文章が与えられた場合には語彙的多様性の指標が高くなった一方で、発話量や流暢さの指標に変化は見られなかった。しかし、実験 2A のデザインでは、RQ2A-3 のデザインでは、正しく刺激文の提示方法が発話に与える影響について調査ができなかったと判断したため、実験 2B では

この問題点を解決するための計画を立て、刺激文の長さと提示方法が発話に与える影響について分析を行った。実験 2B では、日本人大学生と高専生 63 名が 4 種類 (リーディング 100 語・リーディング 150 語・リスニング 100 語・リスニング 150 語) の再話課題に解答する計画で実験を実施した。本実験の RQ は下記のとおりである。

RQ2B-1: Does the text length of input materials affect performance in retelling tasks?

RQ2B-2: Does the information presentation mode (reading or listening) affect performance in retelling tasks?

実験の結果、刺激文がリーディング形式で与えられた場合にはリスニング形式で与えられた場合に比べて発話が長くなる傾向があり、全体的にはリスニング形式で刺激が与えられた場合にはパフォーマンスが低下する傾向があった。ただし、流暢さの指標のうちポーズの合計長に限れば、リスニング 100 語の条件で流暢さが最も高くなったと言える。この結果から、刺激文の提示がリスニング形式で行われる再話課題は発話を多く引き出せなかったため、難易度が高いと想定された。そのため、刺激をリスニング形式で与える場合、学習者・受験者のレベルを良く考えた上で実施する必要があると結論づけた。

研究 3 では、準備時間 (実験 3A) と再話活動前の音読活動 (実験 3B) が再話課題における発話に与える影響 (実験 3B) が検証された。実験 3A では下記の 2 つの RQ を設定し、再話課題における準備時間の有無が発話に与える影響を調査した。

RQ3A-1: Can test-takers produce more words when they have one minute of preparation time than when they do not have any preparation time?

RQ3A-2: Are test-takers' spoken performances—in terms of complexity, accuracy, and fluency—improved when they have one minute of preparation time?

実験 3A には日本人英語学習者 29 名が参加し、そのうち 21 名のデータが分析対象となった。その結果、Syllables と SPS に関しては準備時間ありでタスクに回答した場合に有意に高

く、1 秒以上のポーズ数も準備時間ありのほうが少なかった。研究 1 や 2 の結果と異なり、発話量に関してはタスク条件間で差が生じなかった。その一方で、これまでと同様に統語的複雑さと統語的正確さにタスク条件間で差は生じなかった。流暢さが向上したという結果は先行研究と同様の結果が見られたとも言えるが、発話の複雑さ・正確さ・流暢さのすべての観点で本研究結果と同様の方向性を示したものはなかった。

実験 3B では、再話課題の回答開始前の操作として、刺激文を音読する活動を取り入れた結果、どの程度パフォーマンスが向上するかを検討することを目的に本実験を行った。本実験を実施した理由は再話の事前タスクとしての音読が発話に与える影響について調査し、事前タスクの有無によってよりパフォーマンスが高まるのであれば、熟達度の低い学習者に再話を使用することが可能となると考えたためである。

RQ3B-1: Can test-takers produce more words when they read the input text aloud?

RQ3B-2: Does test-takers' speech production improve in terms of complexity, accuracy, fluency, and pronunciation when they read the input text aloud?

実験 3B では、日本人大学生 18 名のデータを分析の対象とし、実験協力者を 2 群に分け、実験を行った。実験群 3Ba は最初に Material 3B-2 を 3 分で黙読するよう指示が与えられ、その後すぐに読んだ内容と意見・感想を述べるよう指示を与えられた。一方、実験群 3Bb は同じマテリアルを 2 分で黙読した後、1 度だけ音読をするように指示を与え、その後発話を行うという流れで実験を進めた。その後、タスクを入れ替えてデータを収集した。音読があった場合には、音読なし条件に比べて Types が多く、SPS が高くなった。有意傾向にはとどまったものの、Tokens と Syllables で音読あり条件が多く、WPM が高い傾向が見られた。プレタスクとしての音読が語彙的多様性や統語的エラーに与える影響はないという結果が明らかになった。また、音読活動を行ったにも関わらず、発音面の評価に差が見られなかったことは興味深い結果といえる。音読活動を通じて発音の向上が見られることは想定に難くないが、実験協力者は発話中に発音面に注意を向けることが出来なかったと考えられる。

研究 1～3 の結果から、再話課題における発話のうち、タスク条件の操作によって発話の複雑さと正確さが影響を受けることはなかった。しかし、発話量には明らかに影響があったため、学習者・受験者の発話を引き出せる課題を選択できれば、より多くのフィードバックを与える機会が得られることになるため、どのようなタスク条件が学習者・受験者の発話を最も効率的に引き出せるかを検討することは重要である。

研究 4 では、研究 1～3 で実施されたタスクの難易度を比較するため、多相ラッシュモデルによる分析が行われた。

RQ 4-1: How are topic of texts distributed by difficulty?

RQ 4-2: How are task types distributed by difficulty?

この研究では、EBB2 scale (Hirai & Koizumi, 2013) に 0 の評価項目を使用したものを使用したものの、妥当性が低い結果となったことから、実質的には EBB2 scale (Hirai & Koizumi) の評価基準を用いる事となった。まず、再話課題で提示されたテキストの難易度を推定した結果、7 種類あるテキストの難易度は 4 つのレベルに分かれると判明した一方、難易度については、logit 値が -0.15 から 0.16 の範囲に収まるため、大きな差はないと結論付けられた。タスク条件の難易度は、リーディング × 100 語程度の刺激文が与えられる条件で最も難易度が低く、逆に難易度が高かった条件は 150 語 × リーディングであり、研究 2 で難易度が高い課題であると想定されたリスニング形式で刺激を与える課題は、8 種類のタスク条件の中で 3 番目と 4 番目に難しいという結果になった。この結果の齟齬は、研究 2 では主に CAF の指標を元に解釈が進められた一方、研究 4 では評価者による評定結果が分析の対象であったことに原因があると考えられる。この分析の結果、受験者の平均能力値は 0 を下回っており、タスク・テキストの難易度に比べ、受験者の能力が低かったことから、全体としてタスク条件・刺激の内容に関わらず再話課題は今回の実験参加者にとっては難易度が高い課題であると判断された。

本博士論文で実施された実験の結果から、再話のタスク条件を変更することで、発話量や

流暢さを高めることができた一方、発話の正確さと複雑さはタスク条件を変えた場合でも影響は生じないことが明らかになった。再話は提示する文章の表現や内容を変更し、発話を限定できることから、教育上有効なタスクであることは疑いないが、学習者・受験者に多くのフィードバックを与えるためにも多くの発話を引き出すことができるように、タスクを設定しなければならない。